IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Application No.:	09/706,227
Filing Date:	November 3, 2000
Applicant:	Erling W. Wold, et al.
Group Art Unit:	2626
Examiner:	Michael Opsasnick
Title:	METHOD AND APPARATUS FOR CREATING A UNIQUE AUDIO SIGNATURE

BRIEF ON APPEAL ON BEHALF OF APPELLANTS UNDER 37 C.F.R. §41.37

Mail Stop Appeal Brief – Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

TABLE OF CONTENTS

Page

BRIEF ON APPEAL ON BEHALF OF APPELLANTS

In support of the Notice of Appeal filed on December 17, 2007 appealing the Examiner's Final Rejection of claims 45-60, mailed September 17, 2007, Appellants hereby provide the following Appeal Brief. This is an amended Appeal Brief in response to a Notification of Non-Compliant Appeal Brief dated February 27, 2008.

I. REAL PARTY IN INTEREST

The present application was assigned from the inventors, Erling H. Wold, Thomas L. Blum, Douglas H. Keislar, and James A. Wheaton to Audible Magic Corporation, recorded on November 15, 2000 at reel/frame 011650/0791. Thus, Audible Magic Corporation is the real party in interest.

II. RELATED APPEALS AND INTERFERENCES

The undersigned, the Assignee, and the Appellants do not know of any appeals or interferences which would directly affect or which would be directly affected by, or have a bearing on, the Board's decision in this Appeal.

III. STATUS OF THE CLAIMS

Claims 45-60 are reproduced in the attached Appendix A and are the claims on Appeal. Each of these claims is currently pending in the application.

IV. STATUS OF AMENDMENTS

No amendments have been filed subsequent to the final Office Action.

V. SUMMARY OF THE CLAIMED SUBJECT MATTER

This invention involves a method of determining the identity of a sampled work (Claim 45) and an apparatus implementing that method through circuitry (Claim 53). The method steps in Claim 45 are implemented as circuitry for performing each method step in Claim 53. Thus, the two independent Claims (45 and 53) are functionally identical, so the remainder of this Summary shall describe the method claimed in Claim 45, and that description shall be incorporated by reference for Apparatus Claim 53. Note that in this section, (*P x LL aa-bb*) means (*Page x Lines aa-bb*).

Independent Claim 45 of the the presently claimed invention claims a method for determing the identity of a sampled work comprising: receiving data of a sampled work (FIG. 1, #100; P 9 LL 6-10); segmenting said data of said sampled work into a plurality of segments (FIG. 1, #102; P 9 L 12 through P 10 L 6), said segments having predetermined a predetermined segment size and a predetermined hop size (FIG. 4; P 16 LL 6-16); creating a plurality of signatures (FIG. 1, #104, #106; P 10 LL 8-13; FIGs. 5-6; P 16 L 4 through P 19 L 2) wherein each of plurality of signatures is a signature of one of said plurality of segments and wherein each of said plurality of signatures is of said predetermined segment size and said predetermined hop size (FIG. 4; P 16 LL 6-16); comparing said plurality of signatures of said sampled work to a plurality of reference signatures (P 15 L 10 through P 16 L 2) of each of a plurality of reference works wherein said plurality of reference signatures of each of said plurality of reference works are created from a plurality of segments of said each of said plurality of reference works having a known segment size and a known hop size (FIG. 3; P 15 LL 11-16) and said predetermined hop size of each of said plurality of segment of said sampled work is less than said known

hop size (P 16 L 6 through P 17 L 15); and determining an identity of said sample work responsive to said comparison of said plurality of signatures of said sampled work to said signatures of said plurality of reference works (FIG. 7; P 20 L 27 through P 21 L 4).

Independent claim 53 is an apparatus (FIG. 1 #220) claim implementing the method steps in Claim 45 through circuitry and the description of Claim 45 above is incorporated herein by reference.

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Appellants seek the Board's review of the rejection of Claims 45-60 under 35 U.S.C. § 103(a) as being unpatentable over Schulze (4,918,730) in view of Blum, et al. (5,918,223) in view of Savic, et al. (5,327,521) in further view of Voran (6,092,040).

VII. ARGUMENT

Rejection of Claims under 35 U.S.C. §103(a) in view of Schulze, Blum, Savic, and Voran

In the Office Action dated September 17, 2007, the Examiner rejected Claims 45-60 under 35 U.S.C. § 103(a) as being unpatentable over Schulze (4,918,730) in view of Blum, et al. (5,918,223) in view of Savic, et al. (5,327,521) in further view of Voran (6,092,040). In order to maintain a rejection the Examiner has the burden of providing evidence of prima facie obviousness. See MPEP §2143. See also *In re Vaeck*, 947 F.2d 488, 20 USPQ2d 1438 (Fed. Cir. 1991). In order to prove prima facie obviousness, the Examiner must provide evidence in the prior art of a motivation to combine or modify a reference, a reasonable expectation of success, and a teaching of each and every claimed element. *Id.* The Examiner has failed to provide a teaching of each and every claimed element of the claims.

Note, for the remainder of the Arguments section, a prior art reference of the form (*C x LL aa-bb*) means (*Column X Lines aa-bb*).

Claim 45 claims:

A method for determining an identity of a sampled work, said method comprising:

receiving data of a sampled work;

segmenting said data of said sampled work into a plurality of segments, said segments having predetermined a predetermined segment size and a predetermined hop size;

creating a plurality of signatures wherein each of plurality of signatures is a signature of one of said plurality of segments and wherein each of

said plurality of signatures is of said predetermined segment size and said predetermined hop size;

comparing said plurality of signatures of said sampled work to a plurality of reference signatures of each of a plurality of reference works wherein said plurality of reference signatures of each of said plurality of reference works are created from a plurality of segments of said each of said plurality of reference works having a known segment size and a known hop size and said predetermined hop size of each of said plurality of segment of said sampled work is less than said known hop size; and

determining an identity of said sample work responsive to said comparison of said plurality of signatures of said sampled work to said signatures of said plurality of reference works.

The present claimed invention thus (generally): samples a signal; segments the samples with predetermined segment and hop sizes; creates a signature for each segment; compares the sample segments against signatures for a number of segments of a plurality of references, with the hop size of the sample segments being less than the hop size of the reference segments; and then determining the identity of the sampled work through this comparison.

Of particular note here is the underlined portion of the second to last paragraph that (generally) requires that the segment and hop sizes of the sample signal segments be less than the segment and hop sizes of the segments for the various reference works.

The examiner first suggests that a combination of Schulze, Blum, and Savic teaches using different hop sizes, but to be used on subsections of a single signal. However, the first two references do not show hop sizes at all, and Savic shows using a fixed

segment and hop sizes (C 4, LL 36-43, FIG. 4) of 256 samples sampled at 10 kHz, with a shift distance (i.e. hop size) of 64 samples. Thus, the element of different hop sizes is missing from the Savic reference, and the element of hop sizes is totally missing from the Schulze and Blum references.

Note also that Savic does not involve matching a signal against a database of signals to determine if they are the same, but rather it matches utterances from source and target speakers (C 2, LL 18-22) where the content is known to be the same. The speaker is known, and the issues addressed by the invention are that a person does not repeat himself exactly and there would likely be gaps in the source code book. This is a very different problem than is faced by the present invention and Schulze, and there is no reason to believe that the solutions would be analogous, or would even work between Savic and the present invention. Because of the different problems being addressed, and that the two problems would most likely necessitate very different solutions, it is clear that proscribed hindsight was used to select Savic.

Also, the justification or motivation for combining Savic was to "*overlap… signal processing because it would advantageously produce a smooth spectrum*". This makes engineering sense with Savic, since it's purpose is to match very short segments of a speaker's voice, such as single syllables or even shorter utterances, to a relatively small number of short segments of the same voice in the source code book, and then to substitute a corresponding short segment from the target code book for the closest matching utterance (see Summary, C 2 LL 3-40). There are only a small number of potential matches, since a learning step is utilized to create the source and target code books (C 2 LL

18-34). Also note that the smooth spectrum is created by transforming segments from the time domain into the frequency domain (C 2 LL 23-26) by a Fourier transform (C 2 LL 41-46). The smoothing is to overcome and compensate for the fact that when someone says something a second time, it is highly unlikely to be said identically, and the teaching process is unlikely to cover all possible sounds made by a speaker. But the match does not have to be identical, since a best fit match yields sufficiently good results.

But the examiner does not explain why creating a smooth spectrum would be advantageous with Schulze or with the present invention. Indeed, smoothing does not make sense with either the present invention, nor with the Schulze, which are both aimed at matching samples to a substantially identical reference. There is no reason to believe that smoothing the spectrum of signals would be advantageous at all in attempting to match a sample with a reference work, as is claimed. Indeed, since Schulze and the present invention involve attempting to match samples with reference signals that are substantially the same, and often almost identical, such a smoothing would likely reduce the ability to match since the smoothing would likely often reduce promient features in both the sample and the reference segments being compared that might otherwise aid in identification of a reference work. Additionally, the smoothing comes at a significant cost in resources. Reducing the cost in resources is important in the present invention and the first two references due to their need to match source signal segments against multiple segments of many target works. Thus, Savic is counterproductive, both from resource usage and a degredation in ability to match source and reference signals. For all these reasons, the

previous references teach away from the addition of Savic to the combination of references.

The examiner suggests that the first three references teach differing hop sizes. But the cited portion of Savic (C 7 LL 3-15) describes a variable window or segment size. The next sententence describes a fixed hop size: "*To be sure that S is small enough to avoid time aliasing for the preferred embodiment, a shift length of 64 samples has been chosen*" (C 7 LL 18), with "*The window L in the preferred embodiment is a 256 point speech segment having 10,000 samples per second*" (C 7 LL 10-12). Thus, the Savic hop size is a fixed 64/10,000 of a second.

Savic does not teach the limitation of the hop size for reference segments is different from the hop size of the sampled work. Savic does teach the concept of a hop size (C 4, LL 40-44) where the sample data is segmented into blocks of 256 samples each with a shift distance of 1/4 or 64 samples. There is nothing in Savic that teaches a first or reference set of data has a first hop size and a second or sample set of data that has a second hop size. The purpose of the differing hop sizes in the claims is to provide a greater chance of a two segments matching up with respect to the data in the sample in order to get a greater success in matching a segment of the reference work signature to a signature of a sample. Savic does not have this concern because Savic is teaching a speech transforming system in which a recording of a voice is split into overlapping segments for analysis (C 6 LL 7-19 and C 7 LL 3-17). Since there are no comparisons of segments of different items, Savic cannot teach that the hop sizes of the two items are different to try to make matches more likely.

The examiner then suggests that different hop sizes be used with Voran. As noted above, the examiner failed to show diferent hop sizes in the first three references. Furthermore, Voran is not analogous art. It does not involve comparing a sample of a signal to a database of reference samples to determine whether a match can be found, but rather, a first signal against a second signal to determine the time lag between them (Abstract). Note also that in Voran, the two signals are known (or at least presumed) to match closely. That is just the opposite of the present invention which is attempting to find a match. The two problems being solved are radically different, and therefore, it is apparent that proscribed hindsight is again being utilized to combine this reference.

Furthermore, the examiner failed to show where different hop sizes are shown or taught in Voran. A three column section of Voran (C 8 L 44 through C 11 L 44) is hardly the level of specificity required by the MPEP (e.g. § 706.02(j)) or case law. This section appears to involve fixed sampling rates and sample sizes.

Voran does not teach that the segments from the two different sets of data have different hop sizes. A hop size is the amount of data between the start of adjacent segments of data. In claim 45, the data from the unknown work is divided into segments having a first hop size between samples and signatures are generated of each segment. These signatures of these segments are compared to signatures of a segments of a second reference having a second hop size. This allows for a greater possibility that the reference and test segments will align producing better results. Voran does not teach this concept. Instead, Voran teaches an algorithm that detects delay between an input and an output signal (C 2 LL12-24). A second algorithm may be performed to determine the diffe-

rence in speech quality of two signals. Neither of these use differing hop sizes to compare the two signals. In fact there is no mention anywhere in the entirety of Voran about hop sizes between segments. The only mention of the actual samples is given in block 110 and described on Col. 5, LL 32-39 in which Voran states the same number of samples from each signal are stored in separate arrays. The only other mention of the samples is Col. 12, LL 24-32 in which Voran states the reference files must have the same length and be synchronized. This means that both files must have the same number of segments or samples since a different hop size between segments would create more segments and files could not be synchronized. Thus, not only does Voran not teach the use of hop sizes, Voran implicitly teaches against their use. Therefore, Voran does not teach the use of different hop sizes in the data of the two different works.

Furthermore, the claim actually claims a sample signal segment hop size smaller than the hop size of the reference signal segments. That is not the same or even equivalent of Voran teaching "*using different hop sizes when comparing test signals to reference signals*" (OA ¶ 4).

Furthermore, it can be seen from Voran FIG. 2 that that invention would not work at all with the problem addressed by the currently claimed invention. The cost of the Speech Signal Preparation Algorithm 210, 260 and the Frequency Domain Transformation Algorithm 220, 270 would be cost prohibitive in view of attempting to match a sampled signature to a large set of stored signatures for a number of reference works, attempting to identify a matching reference work. It works in Voran since the two signals are known to be mostly the same, one just lagged from the other. But Voran utilizes a

Fourier transform (C 5 LL 25-30) and a low pass filter (C 6 LL 3-5), both of which are expensive and the low bandpass filter would likely degrade the performance of the combination of references in respect to any attempt to implement the present invention. Thus, Voran teaches away from being combined with the other three references for this purpose.

Furthermore, the motivation for combining Voran is that "*it would have been an effective way to measure/distinguish speech devices*". Note first that that justification is found in Voran, and thus the other three references cannot teach or suggest that justification. Secondly, it is relevant only for Voran, and has no real relevance to the other three references. And, indeed, the present invention is not directed toward speech recognition, but rather at matching (typically) audio segments of a sample signal to those of segments of reference works. None of the first three references would appear to be improved by Voran, since they are not aimed at determining lag between two signals known to be the same or similar (Abstract). Also none of the first three references involve speech recognition, nor does the present invention. Savic is closest, but it does not attempt to identify the meaning of speech, but rather blindly substitutes sylabels or short segments of one speaker with those of another. Also, as with Savic, significant extra processing in involved in Voran as is a low pass filter, that are not found in the present invention nor in Schulze. These are counterproductive, and thus also teach away from the combination.

Furthermore, four references were combined to attempt to find the present claims obvious. This number of references is indicia that the claimed invention is nonobvious. Similarly, the fact that the later two references were used to attempt to find small ele-

ments of the presently claimed invention, without showing how the references could be combined, is also indicia of nonobviousness. Indeed, since the motivation for combining the later two references actually teach away from the combination, it is clear that the combination was made based on improper hindsight. Finally, there is no showing how the four references could be effectively combined.

The remainder of the independent claims are similar to this claim, but of a different format or type. Additionally, all the dependent claims are dependent on a claim that has these limitations.

Thus, the four combined references do not show, teach, or suggest the cited claim elements. Furthermore, the justification for the combination of both Savic and Voran are improper. Furthermore, the first two references teach away from combining Savic, and the first three references teach away from combining Voran. Applicants therefore submit that the examiner has not made a prima facie case of obviousness, that this rejection is improper, and request that it be withdrawn.

Claims 46-52 are dependent from claim 45. Thus, claims 46-52 are allowable for at least the same reasons as claim 45. Therefore, Applicants respectfully request that claims 46-52 be allowed.

Furthermore, the examiner failed to point out where in the cited references the elements added for claims 46, 47, and 48 were found.

Claim 53 recites an apparatus that performs the method of claim 45. Thus, claim 53 is allowable for at least the same reasons as claim 45. Thus, Applicant respectfully requests that claim 53 be allowed.

Claims 54-60 are dependent from claim 53. Thus, claims 54-60 are allowable for at least the same reasons as claim 53. Therefore, Applicants respectfully request that claims 54-60 be allowed.

CONCLUSION

Appellants respectfully request the Honorable Board of Patent Appeals and Interferences to reverse the Examiner's rejection of Claims 45-60 under 35 U.S.C. § 103(a) as being unpatentable over Schulze (4,918,730) in view of Blum, et al. (5,918,223) in view of Savic, et al. (5,327,521) in further view of Voran (6,092,040).

Appellants respectfully submit that the prior art does not teach all of the steps performed by a client and that the combination of references is taught away from and is likely inoperable. Accordingly, for at least the aforementioned reasons, Appellants respectfully request the Honorable members of the Board of Patent Appeals and Interferences to reverse the outstanding rejections in connection with the present application and permit each of the claims in connection with the present application be allowed.

Should there be any outstanding matters that need to be resolved in the present application, the Examiner is respectfully requested to contact Bruce E. Hayden, Reg. No. 35,539 at the telephone number of the undersigned below.

If necessary, the Commissioner is hereby authorized in this, concurrent, and future replies, to charge payment or credit any overpayment to Deposit Account No. 50-0612 for any additional fees required under 37 C.F.R. § 1.16 or under 37 C.F.R. § 1.17; particularly, extension of time fees.

Respectfully submitted,
SIERRA PATENT GROUP, LTD

Date: March 4, 2008

By: /Bruce E. Hayden, Reg# 35,539/
Bruce E. Hayden, Reg. No. 35,539

Please address all correspondence to:
Sierra Patent Group, Ltd.
1663 Hwy. 395, Suite 201
Minden, NV 89423
(775) 586-9500
(775) 586-9550

VIII. CLAIMS APPENDIX

This is a complete and current listing of the claims, marked with status identifiers in parentheses. The listing has been reformatted for ease of reading and understanding.

1-44. (Canceled)

45. (Previously Presented) A method for determining an identity of a sampled work, said method comprising:

receiving data of a sampled work;

segmenting said data of said sampled work into a plurality of segments, said segments having predetermined a predetermined segment size and a predetermined hop size;

creating a plurality of signatures wherein each of plurality of signatures is a signature of one of said plurality of segments and wherein each of said plurality of signatures is of said predetermined segment size and said predetermined hop size;

comparing said plurality of signatures of said sampled work to a plurality of reference signatures of each of a plurality of reference works wherein said plurality of reference signatures of each of said plurality of reference works are created from a plurality of segments of said each of said plurality of reference works having a known segment size and a known hop size and said predetermined hop size of each of said plurality of segment of said sampled work is less than said known hop size; and

determining an identity of said sample work responsive to said comparison of said plurality of signatures of said sampled work to said signatures of said plurality of reference works.

46. (Previously Presented) The method of claim 45, wherein said act of creating a signature of said sampled work compromises calculating segment feature vectors for each segment of said sampled work.

47. (Previously Presented) The method of claim 45, wherein said act of creating a signature includes calculating a plurality of MFCCs for each said segment.

48. (Previously Presented) The method of claim 45, wherein said act of creating a signature includes calculating a plurality of acoustical features from the group consisting of at least one of loudness, pitch, brightness, bandwidth, spectrum and MFCC coefficients for each said segment.

49. (Previously Presented) The method of claim 45, wherein said sampled work signature comprises a plurality of segments and an identification portion.

50. (Previously Presented) The method of claim 45, wherein said plurality of segments of said sampled work signature comprise a segment size of approximately 0.5 to 3 seconds.

51. (Previously Presented) The method of claim 50, wherein said plurality of segments of said sampled work signature comprise a hop size of less than 50% of the segment size.

52. (Previously Presented) The method of claim 50, wherein said plurality of segments of said sampled work signature comprise a hop size of approximately 0.1 seconds.

53. (Previously Presented) An apparatus that determines an identity of a sampled work, said apparatus comprising:

creating a plurality of signatures wherein each of plurality of signatures is a signature of one of said plurality of segments;

comparing said plurality of signatures of said sampled work to a plurality of reference signatures of each of a plurality of reference works wherein said plurality of reference signatures of each of said plurality of reference works are created from a plurality of segments of said each of said plurality of reference works having a known segment size and a known hop size and said predetermined hop size of each of said plurality of segment of said sampled work is less than said known hop size

circuitry configured to receive data of a sampled work;

circuitry configured to segment said data of said sampled work into a plurality of segments wherein each of said segments has predetermined segment size and a predetermined hop size;

circuitry configured to create a plurality of signatures of said sampled work based upon said plurality of segments and wherein each of said plurality of signatures is of said predetermined segment size and said predetermined hop size;

circuitry configured to compare said plurality of signatures of said sampled work to a plurality of signatures of for each of a plurality of reference works created from a plurality of sample segments of each of said plurality of reference works, each of said plurality of signatures of each of said plurality of reference works having a known segment size and a known hop size wherein said predetermined hop size of said each of said plurality of signatures of said sampled work is less than said known hop size; and

circuitry configured to determine said sampled work is one of said reference works based upon said comparison.

54. (Previously Presented) The apparatus of claim 53, wherein said circuitry configured to create a signature of said sampled work compromises circuitry configured to calculate segment feature vectors for each of said plurality of segments of said sampled work.

55. (Previously Presented) The apparatus of claim 53, wherein said circuitry configured to create a signature includes calculating a plurality of MFCCs for each said segment.

56. (Previously Presented) The apparatus of claim 53, wherein said circuitry configured to create a signature includes circuitry configured to calculate one of plurality of acoustical features selected from a group consisting of loudness, pitch, brightness, bandwidth, spectrum and MFCC coefficients for each of said plurality of segments of said sampled works.

57. (Previously Presented) The apparatus of claim 53, wherein said sampled work signature comprises a plurality of segments and an identification portion.

58. (Previously Presented) The apparatus of claim 53, wherein said plurality of segments of said sampled work comprise said predetermined segment size of approximately 0.5 to 3 seconds.

59. (Previously Presented) The apparatus of claim 58, wherein said predetermined hop size of said plurality of segments of said sampled work signature is less than 50% of the segment size.

60. (Previously Presented) The apparatus of claim 58, wherein said predetermined hop size of each of said plurality of segments of said sampled work signature is approximately 0.1 seconds.

IX. EVIDENCE APPENDIX

Evidence limited to cited cases.

X. RELATED PROCEEDINGS APPENDIX.

None